

# Correct Convergence of Min-Sum Loopy Belief Propagation in a Block Interpolation Problem

Yutong Wang, Matthew G. Reyes, David L. Neuhoff

February 22, 2017

## Abstract

This work proves a new result on the correct convergence of Min-Sum Loopy Belief Propagation (LBP) in an interpolation problem on a square grid graph. The focus is on the notion of local solutions, a numerical quantity attached to each site of the graph that can be used for obtaining MAP estimates. The main result is that over an  $N \times N$  grid graph with a one-run boundary configuration, the local solutions at each  $i \in B$  can be calculated using Min-Sum LBP by passing difference messages in  $2N$  iterations, which parallels the well-known convergence time in trees.

## 1 Introduction

his paper demonstrates the correct convergence of Loopy Belief Propagation (LBP) in the MAP interpolation of a block of sites given a configuration on its boundary, in the context of a uniform Ising Markov random field. There has been considerable work in analyzing the performance of LBP in the context of maximization problems, for example [1–4]. This paper presents both a new problem setting for Belief Propagation (BP) and a new method of analysis.

In the context of Markov models, a very natural setting is that of MAP estimation of a subset of sites conditioned on a configuration on its boundary, as the Markov property itself tells us that a subset of sites is conditionally independent of all other sites if we know the configuration on the subset’s boundary. Markov models are often expressed as products of functions on single nodes and edges of the associated Markov graph. Thus by taking the negative logarithm of the probability, MAP estimation can be formulated as what is referred as a *min-sum* problem, that of finding configurations that minimize a sum of functions defined on single nodes and edges of the graph. Belief Propagation is a

---

An abbreviated version of this paper has been submitted to ISIT 2017.

Authors addresses: Yutong Wang, Matthew G. Reyes, David L. Neuhoff, EECS Dept., University of Michigan; email: {yutongw, mgreyes, neuhoff}@umich.edu.

recursive distributed algorithm that can be applied to a min-sum problem.<sup>1</sup>

The Markov model considered in this paper is a uniform Ising model with positive correlation on a square grid graph with edges connecting horizontally and vertically adjacent nodes and with nodes assigned values  $+1$  (black) or  $-1$  (white) [5]. This is a single-parameter binary model that favors configurations in which neighboring nodes have the same value. Edges on which the two endpoints have different values are called *odd bonds*. Our problem is to MAP estimate the configuration  $\mathbf{x}_B$  on a subset  $B$  of sites conditioned on a boundary configuration  $\mathbf{x}_{\partial B}$ . In this context, MAP estimation amounts to finding configurations that minimize the sum

$$O(\mathbf{x}_B, \mathbf{x}_{\partial B}) = \sum_{\{i,j\}: i \in B} \mathbb{I}_{(x_j \neq x_i)} \quad (1)$$

of odd bonds over all edges in the graph with at least one endpoint in  $B$ . This problem arose in the context of an image compression application modeling binary images as instances of such an Ising model [6] and also in the context of grayscale image reconstruction [7]. Analytical solutions for the set of MAP configurations conditioned on boundary configurations containing 2 or 4 odd bonds have been found [6, 8]. Such boundaries are termed, respectively, *1-run* and *2-run boundaries*. The MAP configurations on a block conditioned on its boundary are referred to as *global solutions* for the boundary.

Min-Sum LBP is a popular distributed message-passing algorithm for minimizing a sum of functions defined on edges of a graph. As a distributed algorithm, it does not attempt to compute global solutions, but rather, for each site, the minimum numbers of odd bonds in configurations where site  $i$  is black ( $x_i = +1$ ) or white ( $x_i = -1$ ),

$$O_i^*(\pm 1, \mathbf{x}_{\partial B}) := \min_{\mathbf{x}_B: x_i = \pm 1} O(\mathbf{x}_B, \mathbf{x}_{\partial B}). \quad (2)$$

These minimum numbers of odd bonds provide some information regarding the set of global solutions. For example, if  $O_i^*(-1, \mathbf{x}_{\partial B}) < O_i^*(1, \mathbf{x}_{\partial B})$ , then we can say that site  $i$  has value  $-1$  in all global solutions, whereas if  $O_i^*(-1, \mathbf{x}_{\partial B}) > O_i^*(1, \mathbf{x}_{\partial B})$ , site  $i$  has value  $1$  in all global solutions. On the other hand, if  $O_i^*(-1, \mathbf{x}_{\partial B}) = O_i^*(1, \mathbf{x}_{\partial B})$ , what we can say is that there exists a global solution in which site  $i$  has value  $-1$ , and there exists a global solution in which site  $i$  has value  $1$ . Moreover, as pointed out in [2] when there are multiple sites such that  $O_i^*(-1, \mathbf{x}_{\partial B}) = O_i^*(1, \mathbf{x}_{\partial B})$ , a joint configuration on these sites cannot be chosen independently of each other.

In practice, the messages are normalized to prevent numerical overflow. As a result, the goal of BP becomes computing the difference

$$o_i^*(\mathbf{x}_{\partial B}) = O_i^*(-1, \mathbf{x}_{\partial B}) - O_i^*(1, \mathbf{x}_{\partial B}), \quad (3)$$

which we refer to as the *local solution* at site  $i$  given boundary configuration  $\mathbf{x}_{\partial B}$ . At the  $n$ -th iteration of message-passing, an estimate  $\hat{o}_i^n(\mathbf{x}_{\partial B})$  of the local

---

<sup>1</sup>Belief Propagation in the context of MAP estimation is more often studied as a max-product problem, the variant obtained without taking the negative logarithm.

solution at site  $i$  is produced. If  $B$  were a tree, i.e., an acyclic graph, the usual argument of the correct convergence of BP on trees could be adapted here to show that  $\hat{o}_i^n(\mathbf{x}_{\partial B})$  converges  $o_i^*(\mathbf{x}_{\partial B})$ . For cyclic graphs such as the grid graphs considered in the present paper, general convergence is unknown except in special cases such as when the graph is a single cycle [1]. However, it was observed in [8] that empirically LBP converged to the correct local solutions for a 1-run boundary.

It is our belief that LBP can be an effective distributed algorithm for the MAP interpolation problem posed here. While a complete understanding of the correct convergence properties of LBP is currently beyond our means, in this paper we prove in Theorem 6 that it correctly converges in the case where  $B$  is an  $N \times N$  grid graph with horizontal and vertical edges with a one-run boundary configuration. Specifically, we show that the local solution at each  $i \in B$  can be calculated using Min-Sum Belief Propagation by passing difference messages in  $2N$  iterations. We define the *Forward* and *Backward Convergence Property*, which are crucial for our analysis of the convergence. To verify the correctness of the converged results of LBP, we use Proposition 4, which is proven by leveraging the results in [6]. Thus, the results of this paper demonstrate that at least in the case of one-run boundaries, LBP converges to the correct local solution in what amounts with a minimal number of iterations. We hope our work here gives some theoretical justification for using LBP and local solutions for interpolation problems beyond this setting.

The remainder of this paper is as organized as follows. In Section 2, we introduce background on graphs, the boundary interpolation problem, and Belief Propagation. In Section 3, we introduce our message recursion, local solutions, and state what the correct local solutions are. In Section 4, we introduce and discuss the concepts of forward and backward convergence used to prove our results. In Section 5, we present the proof of our main result, Theorem 6.

## 2 Background and Problem Formulation

We introduce notation on graphs, configurations, etc. Let  $\mathbb{N} = \{1, 2, \dots\}$  and  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ . Edges in an undirected graph are written as  $\{i, j\}$ . In a directed graph, edges are written as  $j \rightarrow i$ . For an undirected graph  $G = (V, E)$  and a subset  $S \subseteq V$ , we let  $\partial S := \{j \in V \setminus S \mid \exists i \in S : \{j, i\} \in E\}$ . Abusing notation, we often refer to a subset  $S \subseteq V$  as if it is a subgraph. For instance, the statement “suppose  $S$  is connected” means “suppose the  $G$ -induced subgraph on  $S$  is connected”. Another abuse of notation is  $\partial i := \partial\{i\}$ .

### 2.1 Grid graphs, configurations, and odd bonds

In this subsection, we define the setting that we work in for the majority of the paper. Let  $G = (V, E)$  be the  $(N + 2) \times (N + 2)$  *grid graph* with the *4-neighbor topology* in which the sites are arranged in a square lattice and the edges consist of horizontally and vertically adjacent sites of  $V$ . Two sites connected by an

edge are referred to as *neighbors*. The *interior* is the set of sites  $B \subseteq V$  having four neighbors. The set  $\partial B$  is the *boundary* of  $B$ , i.e.,  $\partial B = V \setminus B$ .

For each site  $i \in V$ ,  $x_i \in \{-1, 1\}$  is an assignment to site  $i$ . An assignment to a set of sites  $S$  is called a *configuration* and denoted  $\mathbf{x}_S$ . For concreteness,  $x_i = 1$  (resp.  $x_i = -1$ ) means that site  $i$  is colored black (resp. white). An edge  $\{i, j\}$  with  $x_i \neq x_j$  is called an *odd bond*. A configuration  $\mathbf{x}_{\partial B}$  on  $\partial B$  is called a *boundary configuration*. Finally, we define one-run boundaries:

**Definition 1 (One-run boundary).** Let  $\mathbf{x}_{\partial B}$  be a boundary configuration. Define  $R^+ = \{i \in \partial B \mid x_i = +1\}$  and define  $R^-$  similarly. We say that  $\mathbf{x}_{\partial B}$  is a *one-run configuration* if  $R^+$  is a connected subgraph.

## 2.2 MAP Estimation and Global Solutions

Given  $\mathbf{x}_{\partial B}$  and interior configuration  $\mathbf{x}_B$ , the quantity  $O(\mathbf{x}_B, \mathbf{x}_{\partial B})$  as defined by equation (1) is the number of odd bonds within  $B$  and between  $B$  and  $\partial B$ . The *global minimum* number of odd bonds between an interior configuration  $\mathbf{x}_B$  and the given boundary configuration  $\mathbf{x}_{\partial B}$  is

$$O_B^*(\mathbf{x}_{\partial B}) := \min_{\mathbf{x}_B} O(\mathbf{x}_B, \mathbf{x}_{\partial B}).$$

In [6, 8], all MAP solutions were found for all one-run boundary configurations and at least one MAP solution was found for every two-run boundary configuration. Specifically, for boundaries consisting of one-run of white and one-run of black, the MAP solutions consisted of configurations generated by a shortest path connecting the endpoints of either runs. In this work, we refer to a MAP solution as a *global solution*.

## 2.3 Belief Propagation

We first review BP for interpolating from the boundary in the context of a tree. For now, let  $G = (V, E)$  be an arbitrary tree apart from the grid graph currently under consideration. Let  $B \subseteq V$  be a subtree and  $\mathbf{x}_{\partial B}$  a boundary configuration. Consider any two adjacent nodes  $i, j \in B$ . Removing edge  $\{i, j\}$  from  $B$  disconnects  $B$  into two connected components. Let  $T_j^i \subseteq B$  be the connected component containing  $j$ . Define  $M_{j \rightarrow i}(-1)$  to be the minimal number of odd bonds in  $T_j^i \cup \partial B$  over all possible configurations  $\mathbf{x}_{T_j^i}$  on  $T_j^i$  plus the odd bond between  $j$  and  $i$ , if any. More precisely,

$$M_{j \rightarrow i}(-1) = \mathbb{I}_{(x_j \neq 1)} + \min_{\mathbf{x}_{T_j^i}} O(\mathbf{x}_{T_j^i}, \mathbf{x}_{\partial B}).$$

Recall  $O_i^*(-1, \mathbf{x}_{\partial B})$  as defined by equation (2). Below, we suppress the dependency on  $\mathbf{x}_{\partial B}$  and simply write  $O_i^*(-1)$ . Likewise, we write  $O_i^*(1)$ . Since  $B$  is a tree, it is easy to see that  $O_i^*(-1)$  could be expressed as

$$O_i^*(-1) = \sum_{j \in \partial i} M_{j \rightarrow i}(-1). \quad (4)$$

We define  $M_{j \rightarrow i} \triangleq [M_{j \rightarrow i}(-1), M_{j \rightarrow i}(1)]$  to be the *message* from site  $j$  to site  $i$ . Using a recursive argument, it is straightforward to show

$$M_{j \rightarrow i}(x_i) = \min_{x_j \in \{\pm 1\}} \left\{ \mathbb{I}_{(x_i \neq x_j)} + \sum_{k \in \partial j \setminus i} M_{k \rightarrow j}(x_j) \right\}$$

The above recursion relation induces a message-passing algorithm:

**Definition 2.** Given a boundary configuration  $\mathbf{x}_{\partial B}$ , for each edge  $\{i, j\} \in E$  such that  $i \in B$  or  $j \in B$  define

*Boundary condition:*  $M_{j \rightarrow i}^n(x_i) := \mathbb{I}_{(x_j \neq x_i)}$  for all  $n \geq 0$ , if  $j \in \partial B$ ,

*Initialization:*  $M_{j \rightarrow i}^0(x_i) := 0$  if  $j \in B$ ,

*Update:* If  $j \in B$  and  $n > 0$ , then

$$M_{j \rightarrow i}^n(x_i) := \min_{x_j \in \{\pm 1\}} \left\{ \mathbb{I}_{(x_i \neq x_j)} + \sum_{k \in \partial j \setminus i} M_{k \rightarrow j}^{n-1}(x_j) \right\}$$

Since the graph  $B$  considered in this subsection is acyclic, this algorithm is referred to as Belief Propagation (BP), or more specifically, as Min-Sum BP. After a number of iterations equal to the length of the longest path in  $B$ , equation (4) permits computation of  $O_i^*(\pm 1)$  for each site in  $B$ . For cyclic graphs, such as the graph considered in this paper as defined in Section 2.1, the algorithm above can still be used and is referred to as *Loopy* Belief Propagation.

### 3 Difference messages and local solutions

For the remainder of this paper, we are in the setting of Section 2.1. To avoid numerical overflow, it is standard practice to normalize the messages in LBP. In our case, we pass the *difference messages*  $m_{j \rightarrow i}^n := M_{j \rightarrow i}^n(-1) - M_{j \rightarrow i}^n(1)$  which satisfy the following recursion.

**Lemma 3 (Difference Messages).** *Definition 2 induces the following message-passing dynamics on the difference messages as follows: Given a boundary configuration  $\mathbf{x}_{\partial B}$ , for each edge  $\{i, j\} \in E$  such that  $i \in B$  or  $j \in B$  we have*

*Boundary condition:*  $m_{j \rightarrow i}^n := x_j$  for all  $n \geq 0$ , if  $j \in \partial B$ ,

*Initialization:*  $m_{j \rightarrow i}^0 := 0$  if  $j \in B$ ,

*Update:* If  $j \in B$  and  $n > 0$ , then

$$m_{j \rightarrow i}^n := \text{sign} \left\{ \sum_{k \in \partial j \setminus i} m_{k \rightarrow j}^{n-1} \right\}$$

where  $\text{sign}(t) = t/|t|$  for  $t \neq 0$  and  $\text{sign}(0) = 0$ .

*Proof.* For the *boundary condition*, if  $j \in \partial B$ , then  $m_{j \rightarrow i}^n = M_{j \rightarrow i}^n(-1) - M_{j \rightarrow i}^n(1) = \mathbb{I}(x_j \neq -1) - \mathbb{I}(x_j \neq 1) = x_j$ .

There is nothing to check for the *initialization*.

For the *update*, fix an  $n > 0$  and an edge  $\{i, j\} \in E$  such that  $j \in B$ . For  $z \in \{\pm 1\}$ , let  $\Phi(z) := \sum_{k \in \partial j \setminus i} M_{k \rightarrow j}^{n-1}(z)$ . By definition, we have

$$\begin{aligned} m_{j \rightarrow i}^n &= M_{j \rightarrow i}^n(1) - M_{j \rightarrow i}^n(-1) \\ &= \min_{z \in \{\pm 1\}} \{\mathbb{I}(1 \neq z) + \Phi(z)\} \\ &\quad - \min_{z \in \{\pm 1\}} \{\mathbb{I}(-1 \neq z) + \Phi(z)\} \\ &= \min \{\Phi(1), 1 + \Phi(-1)\} \\ &\quad - \min \{1 + \Phi(1), \Phi(-1)\} \end{aligned}$$

From the above and the fact that  $\Phi_j$  maps into  $\mathbb{Z}$ , one gets

$$\begin{aligned} m_{j \rightarrow i}^n &= \begin{cases} -1 & \Phi(1) > \Phi(-1) \\ 1 & \Phi(1) < \Phi(-1) \\ 0 & \Phi(1) = \Phi(-1) \end{cases} \\ &= \text{sign}\{\Phi(-1) - \Phi(1)\} \end{aligned}$$

By definitions,  $\Phi(-1) - \Phi(1) = \sum_{k \in \partial j \setminus i} M_{k \rightarrow j}^{n-1}(-1) - M_{k \rightarrow j}^{n-1}(1) = \sum_{k \in \partial j \setminus i} m_{k \rightarrow j}^{n-1}$ . Thus, we are done.  $\square$

By passing difference messages  $\{m_{i \rightarrow j}^n\}$  rather than the original messages  $\{M_{i \rightarrow j}^n\}$  we are unable to compute the quantities  $O_i^*(1)$  and  $O_i^*(-1)$ . Nevertheless, we can use the difference messages to estimate the local solutions  $o_i^*(\mathbf{x}_{\partial B})$  as defined by equation (3). In Theorem 6, we show that these estimates indeed converge to the truth. Below, we make the dependence of  $o_i^*(\mathbf{x}_{\partial B})$  on the boundary condition  $\mathbf{x}_{\partial B}$  implicit and simply write  $o_i^*$ .

Before discussing the convergence of Min-Sum LBP, we first use the global solutions found in [6, 8] to directly compute the local solutions for sites  $i \in B$ . Using these local solutions, we will in Section 5 show that Min-Sum LBP converges to the correct local solutions for 1-run boundaries.

Let  $\mathbf{x}_{\partial B}$  be a one-run boundary. A *positive simple path* is a subset  $P^+ = \{i_1, \dots, i_k\}$  of  $V$  such that

1. The subgraph  $P^+$  is a path,
2.  $i_1, i_k \in \partial B$  are the two endpoints of  $P^+$ ,
3. for  $1 < \ell < k$ , either  $i_\ell \in B$  or, if  $i_\ell \in \partial B$ ,  $x_{i_\ell} = 1$ ,
4.  $k$  is minimal satisfying the above properties.

A *negative simple path*  $P^-$  is defined in the same way by replacing  $x_{i_\ell} = 1$  with  $x_{i_\ell} = -1$  in item 3 above. Define the *positive inner (resp. outer) path*  $P_I^+$

(resp.  $P_O^+$ ) be the positive simple path that minimizes (resp. maximizes) over all positive simple paths  $P^+$  the number of nodes enclosed by  $P^+ \cup R^+$ . Similarly, define  $P_I^-$  and  $P_O^-$ . See examples in Figure 1.

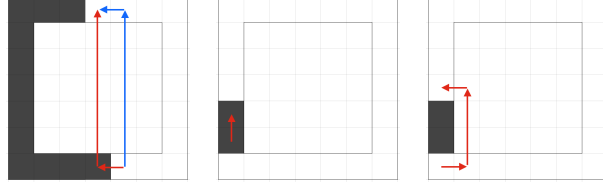


Figure 1: Left: Red path =  $P_O^-$ , blue path =  $P_I^-$ . Middle & right:  $P_O^+ = P_I^+$ .

Define  $I^+$  to be the set  $P_I^+ \cup R^+$  together with the set of nodes enclosed by them. Similarly, define  $O^+$ ,  $I^-$  and  $O^-$ . We use the calligraphic font to denote intersection of these sets with  $B$ , e.g.  $\mathcal{I}^+ = I^+ \cap B$ ,  $\mathcal{O}^+ = O^+ \cap B$  and so on. We call  $\mathcal{I}^+$  the *positive inner region*  $\mathcal{O}^+$  the *positive outer region*, and so on. Define  $\delta\mathcal{I}^+ = \{v \in \mathcal{I}^+ \mid v \text{ touches } O^-\}$  and  $\delta\mathcal{I}^- = \{v \in \mathcal{I}^- \mid v \text{ touches } O^+\}$ . Note that  $\delta\mathcal{I}^+$  is not the same as  $\partial\mathcal{I}^+$ , with which we are not concerned. The following proposition gives the local solutions over regions that we have defined.

**Proposition 4.** *For a one-run boundary  $\mathbf{x}_{\partial B}$ ,*

$$o_i^*(\mathbf{x}_{\partial B}) = \begin{cases} \pm 4 & : i \in \mathcal{I}^\pm \setminus \delta\mathcal{I}^\pm \\ \pm 2 & : i \in \delta\mathcal{I}^\pm \\ 0 & : i \in \mathcal{O}^+ \cap \mathcal{O}^- \end{cases}$$

*are the local solutions for sites  $i \in B$ .*

*Proof.* Let  $i \in \mathcal{O}^+ \cap \mathcal{O}^-$ . Thus  $i \in \mathcal{O}^+$ , which means that there is a MAP configuration, namely the one generated by the black outer path, in which site  $i$  is black. Likewise,  $i \in \mathcal{O}^-$ , which means that there is a MAP configuration in which site  $i$  is white. Therefore,  $O_i^*(-1) = O_i^*(1)$ , and hence  $o_i^* = 0$ .

Now let  $i \in \mathcal{I}^+ \setminus \delta\mathcal{I}^+$ . Since  $i \in \mathcal{I}^+$ , site  $i$  is black in all MAP configurations, therefore  $O_i^*(1) < O_i^*(-1)$ . Moreover, since  $i \notin \delta\mathcal{I}^+$ , in all MAP configurations, every neighbor  $j \in \partial i$  is also black. Therefore, in any MAP configuration, if we flip site  $i$  to white, then, this will increase the number of odd bonds by 4. If we also flip a neighbor  $j \in \partial i$  to white, this will further increase the number of odd bonds by at least 2. Therefore  $O_i^*(-1) = O_i^*(1) + 4$ .

Let  $i \in \delta\mathcal{I}^+$ . Since  $i \in \mathcal{I}^+$ , site  $i$  is black in all MAP configurations, therefore  $O_i^*(1) < O_i^*(-1)$ . By definition of  $\delta\mathcal{I}^+$ , there is a MAP configuration  $x_B$  in which a neighbor  $j \in \partial i$  is white. We claim that there can only be one such neighbor  $j$ . This is because if there were two such neighbors, then we could flip  $i$  to white and keep the number of odd bonds the same, which would imply that there is a MAP configuration in which  $i$  is white. If there were more than two such neighbors  $j \in \partial i$ , then flipping  $i$  to white would strictly

decrease the number of odd bonds, which contradicts  $O_i^*(1) < O_i^*(-1)$ . Thus,  $O_i^*(-1) = O_i^*(1) + 2$ .

The remaining two cases follow from arguments analogous to those in the previous two paragraphs. This completes the proof.  $\square$

For each  $i \in B$ , define  $\hat{o}_i^n := \sum_{j \in \partial i} m_{j \rightarrow i}^n$  to be the *estimates* of the local solution  $o_i^*$ . Applying the usual arguments for correct convergence of BP on trees (see [9] for instance) gives the following proposition

**Proposition 5.** *Let  $G = (V, E)$  be a tree and  $B \subseteq V$  be a subtree, and  $\mathbf{x}_{\partial B}$  be any boundary configuration. Then the difference messages  $m_{j \rightarrow i}^n$  converge and the estimates  $\hat{o}_i^n$  converge to  $o_i^*$  in number of steps equal to the diameter of the tree plus 1.*

We now present our main result.

**Theorem 6.** *Let  $G = (V, E)$  and  $B \subseteq V$  be as defined in Section 2.1, and  $\mathbf{x}_{\partial B}$  be a one-run boundary configuration. Then for every edge  $\{j, i\}$  in  $G$  with  $j \in B$ , the difference messages  $m_{j \rightarrow i}^n$  converge and the estimates  $\hat{o}_i^n$  converge to  $o_i^*$  in number of steps equal to  $2N$ .*

We note that the diameter of  $B$  is  $2N - 2$  rather than  $2N$ . Thus, Theorem 6 parallels Proposition 5.

## 4 Forward and backward convergence

In this section, we present the two technical lemmas, the Forward and Backward Convergence Lemmas, that allow us to prove the convergence of the difference messages in the setting of Section 2.1. The main idea is that convergence of messages on certain rectangular subsets of  $B$  takes place in two phases, forward from the corners of the graph, and then backward towards the boundary.

Recall that we are in the setting of Section 2.1. We identify  $V$  with the point set  $\{0, 1, \dots, N\}^2$  in the usual way, i.e.,  $(0, 0)$  is the bottom-left corner,  $(0, N+1)$  is the top-left corner, and so on. We represent nodes in  $V$  in two ways. The first way uses  $i, j, k, l, m$  (with possible subscripts) to represent a vertex in a coordinate-free way, e.g.,  $i_1, i_2 \in V$ . The second way uses  $(a, b), (\alpha, \beta)$  (again, with possible subscripts) to represent a vertex in coordinates, e.g.  $(a, b) \in V$  where  $a, b \in \{0, 1, \dots, N+1\}$ . At times, we will say “let  $i = (a, b) \in V$  be a vertex” to simultaneously refer to both representations.

Define  $\vec{G} = (V, \vec{E})$  to be the directed graph such that the vertex set of  $\vec{G}$  and  $G = (V, E)$  are the same, and for each undirected edge  $\{i, j\} \in E$ , there are exactly two directed edges  $i \rightarrow j$  and  $j \rightarrow i$  in  $\vec{E}$ . Let  $\{\mathcal{N}, \mathcal{S}, \mathcal{E}, \mathcal{W}\}$  be the set of *directions* north, south, east, and west, respectively. Define corresponding direction vectors  $v(\mathcal{N}) = (0, 1)$ ,  $v(\mathcal{S}) = (0, -1)$ ,  $v(\mathcal{E}) = (1, 0)$  and  $v(\mathcal{W}) = (-1, 0)$ . For each  $(a, b) \in V$  and direction  $\mathcal{D}$ , if  $(a, b) + v(\mathcal{D}) \in V$ , where  $+$  is just the usual vector summation, then  $(a, b) + v(\mathcal{D})$  is said to be the  $\mathcal{D}$  *neighbor* of  $(a, b)$ . For example,  $(a+1, b)$  is the eastern neighbor of  $(a, b)$ . Given a node



$(a, b) \in V$  and a direction  $\mathcal{D}$  such that  $(a, b) + v(\mathcal{D}) \in V$ , define  $\mathcal{D}(a, b) \in \vec{E}$  as  $\mathcal{D}(a, b) := (a, b) \rightarrow (a, b) + v(\mathcal{D})$ . For example,  $\mathcal{E}(a, b) = (a, b) \rightarrow (a + 1, b)$ .

For a message  $m_{j \rightarrow i}^n$  as defined in Lemma 3, we often drop the superscript and simply write  $m_{j \rightarrow i}$  when the time index is not of concern. With this notation,  $m_{\mathcal{E}(a, b)}$  denotes the message from  $(a, b)$  to its eastern neighbors and so on. Now, given a subset  $S \subseteq B$  and a direction  $\mathcal{D}$ , a message *received by  $S$  from direction  $\mathcal{D}$*  is a message of the form  $m_{\mathcal{D}(a, b)}$  for some  $(a, b) \notin S$  such that  $(a, b) + v(\mathcal{D}) \in S$ . A message *sent from  $S$  in the direction  $\mathcal{D}$*  is a message of the form  $m_{\mathcal{D}(a, b)}$  for some  $(a, b) \in S$ . Note that according to these definition, a message  $m_{\mathcal{D}(a, b)}$  sent from  $S$  is not considered to be received by  $S$  even if  $(a, b) + v(\mathcal{D})$  is in  $S$ . See Figure 3 which illustrates the messages sent from and received by  $S$  where  $S$  is the set of blue nodes.

The sets  $S$  for which we are interested in messages received by and sent from are rectangular subsets of  $B$  defined by the boundary runs. Let  $i_1 = (\alpha_1, \beta_1)$  and  $i_2 = (\alpha_2, \beta_2) \in B$ . Define the *rectangle*  $R_{i_1}^{i_2}$  with corners at  $i_1$  and  $i_2$  by

$$R_{i_1}^{i_2} := \{(a, b) \in B : \min(\alpha_1, \alpha_2) \leq a \leq \max(\alpha_1, \alpha_2), \\ \min(\beta_1, \beta_2) \leq b \leq \max(\beta_1, \beta_2)\}.$$

For  $D \in \mathbb{N}$ , define the *cut-rectangle*  $R_{i_1}^{i_2}(D)$  of nodes of distance  $D - 1$  from the corner  $i_1$  as

$$R_{i_1}^{i_2}(D) := \{(a, b) \in R_{i_1}^{i_2} : |a - \alpha_1| + |b - \beta_1| \leq D - 1\}.$$

Define the *L-shaped region*  $L_{i_1}^{i_2}$  with corner at  $i_1$  to be

$$L_{i_1}^{i_2} := \{(\alpha_1, b) \in B : \min(\beta_1, \beta_2) \leq b \leq \max(\beta_1, \beta_2)\} \\ \cup \{(a, \beta_1) \in B : \min(\alpha_1, \alpha_2) \leq a \leq \max(\alpha_1, \alpha_2)\}.$$

See Figure 2 for examples.

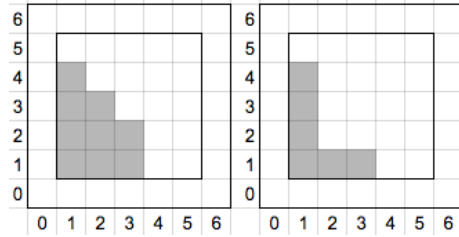


Figure 2: Left:  $R_{(1,1)}^{(3,4)}(4)$ . Right:  $L_{(1,1)}^{(3,4)}$ .

We say the *unordered* pair of directions  $(\mathcal{D}_1, \mathcal{D}_2)$  is *adjacent* if  $v(\mathcal{D}_1)$  and  $v(\mathcal{D}_2)$  are orthogonal. The set of adjacent pairs of directions is

$$\mathcal{A} := \{(\mathcal{E}, \mathcal{N}), (\mathcal{E}, \mathcal{S}), (\mathcal{W}, \mathcal{N}), (\mathcal{W}, \mathcal{S})\}.$$

For two elements  $i_1, i_2 \in B$ , we use the notation  $i_1 \triangleright i_2$  to denote an *ordered* pair of vertices which are not necessarily neighbors.

**Definition 7 (Compatible tuples).** Let  $i_1 = (\alpha_1, \beta_1), i_2 = (\alpha_2, \beta_2) \in B$ . We say that the tuple  $(i_1 \triangleright i_2, \mathcal{D}_1, \mathcal{D}_2)$  is *compatible* if  $(\mathcal{D}_1, \mathcal{D}_2) \in \mathcal{A}$  and there exists non-negative coefficients  $c_1, c_2$  such that  $c_1 v(\mathcal{D}_1) + c_2 v(\mathcal{D}_2) = (\alpha_2 - \alpha_1, \beta_2 - \beta_1)$ .

The ordering of  $i_1 \triangleright i_2$  is crucial because if  $i_1 \neq i_2$ , then  $(i_1 \triangleright i_2, \mathcal{D}_1, \mathcal{D}_2)$  is compatible implies  $(i_2 \triangleright i_1, \mathcal{D}_1, \mathcal{D}_2)$  is *not* compatible. However, the ordering of  $\mathcal{D}_1$  and  $\mathcal{D}_2$  is irrelevant.

**Definition 8 (Convergence of messages).** Let  $\sigma = \pm 1$  and  $N \in \mathbb{N}_0$ . For a given directed edge  $j \rightarrow i \in \vec{E}$ , we say that the messages  $m_{j \rightarrow i}^n$  converges in  $N$  iterations to  $\sigma$  if  $m_{j \rightarrow i}^n = \sigma$  for all  $n \geq N$ .

The definition below is the salient feature of the one-run boundary configuration underlying the proof of our main result

**Definition 9 (Forward Convergence (FC) Property).** The compatible tuple  $(i_1 \triangleright i_2, \mathcal{D}_1, \mathcal{D}_2)$  is *forward convergent* to  $\sigma = \pm 1$  at time  $n_0 \in \mathbb{N}_0$ , abbreviated as  $\text{FC}(\sigma, n_0)$ , if each message received by  $R_{i_1}^{i_2}$  from directions  $\mathcal{D}_1$  and  $\mathcal{D}_2$  converge in  $n_0$  iterations to  $\sigma$ .

We will refer to the following as the *FC Lemma*.

**Lemma 10 (Forward Convergence).** Let  $\sigma = \pm 1$  and  $n_0 \in \mathbb{N}_0$ . Suppose that  $(i_1 \triangleright i_2, \mathcal{D}_1, \mathcal{D}_2)$  is  $\text{FC}(\sigma, n_0)$ . Then all messages sent from  $R_{i_1}^{i_2}$  in the directions  $\mathcal{D}_1$  and  $\mathcal{D}_2$  converge in  $(n_0 + 2N - 1)$  iterations to  $\sigma$ .

*Proof.* Letting  $D = 2N - 1$  in Lemma 11 (proven next), then  $R_{i_1}^{i_2} = R_{i_1}^{i_2}(D)$ . Thus, we have the desired result.  $\square$

Lemma 10 is essentially a corollary of the lemma below.

**Lemma 11.** Let  $\sigma = \pm 1$  and  $n_0 \in \mathbb{N}_0$ . Suppose  $(i_1 \triangleright i_2, \mathcal{D}_1, \mathcal{D}_2)$  is  $\text{FC}(\sigma, n_0)$ . Then for all  $D \in \mathbb{N}$ , all edges sent from  $R_{i_1}^{i_2}(D)$  in the directions  $\mathcal{D}_1$  and  $\mathcal{D}_2$  converge in  $(n_0 + D)$  iteration to  $\sigma$ .

*Proof.* For  $\ell = 1, 2$ , let  $i_\ell = (\alpha_\ell, \beta_\ell)$  be the coordinate representation of  $i_\ell$ . Since  $V$  is embedded in  $\mathbb{R}^2$ , rotations and reflections on  $\mathbb{R}^2$  that preserve  $V$  induce graph automorphisms on  $G = (V, E)$  that respects the adjacency of directions. Thus, because the message-passing dynamics defined in Lemma 3 is isotropic, we can assume  $\alpha_1 \leq \alpha_2, \beta_1 \leq \beta_2$  and  $\mathcal{D}_1 = \mathcal{E}$  and  $\mathcal{D}_2 = \mathcal{N}$  after applying appropriate rotations and reflections. Relabeling  $V$  using the rule  $(a, b) \mapsto (a - \alpha_1 + 1, b - \beta_1 + 1)$ , we may further assume  $\alpha_1 = \beta_1 = 1$ . Relabeling time, we may assume  $n_0 = 0$ . The dynamics defined in Lemma 3 is invariant under multiplication by  $-1$ , i.e., we may replace every instance of  $m_{j \rightarrow i}^n$  by  $-m_{j \rightarrow i}^n$ . Hence, we may assume  $\sigma = 1$ . Let  $(\alpha, \beta) = (\alpha_2, \beta_2)$ .

Given the reductions in the preceding paragraph, we now have only to prove that for each  $(a, b) \in R_{(1,1)}^{(\alpha, \beta)}(D)$  such that  $a \leq \alpha$  and  $b \leq \beta$ , the messages  $m_{\mathcal{E}(a,b)}^n$  and  $m_{\mathcal{N}(a,b)}^n$  converge in  $D$  iterations to 1. In other words, for all  $n \geq D$

$$m_{\mathcal{E}(a,b)}^n := \text{sign}\{m_{\mathcal{E}(a-1,b)}^{n-1} + m_{\mathcal{N}(a,b-1)}^{n-1} + m_{\mathcal{S}(a,b+1)}^{n-1}\} = 1 \quad (5)$$

$$m_{\mathcal{N}(a,b)}^n := \text{sign}\{m_{\mathcal{E}(a-1,b)}^{n-1} + m_{\mathcal{N}(a,b-1)}^{n-1} + m_{\mathcal{W}(a+1,b)}^{n-1}\} = 1 \quad (6)$$

In order to prove (5) and (6), it suffices to show

$$m_{\mathcal{E}(a-1,b)}^{n-1} = m_{\mathcal{N}(a,b-1)}^{n-1} = 1 \quad (7)$$

because in this case the value of the third term inside the sign in (5) and (6) above cannot influence the message.

The boundary conditions part of the definition of property FC(1, 0) translates to the fact that

$$m_{\mathcal{E}(0,b)}^n = m_{\mathcal{N}(a,0)}^n = 1, \forall n \geq 0 \quad (8)$$

for all  $1 \leq a \leq \alpha$  and  $1 \leq b \leq \beta$ .

We proceed by induction on  $D$ . For the base case,  $D = 1$  and  $R_{(1,1)}^{(\alpha,\beta)}(D) = \{(1,1)\}$ . Let  $n \geq 1$ . Observe that (7) follows from (8). Thus, (5) and (6) are proven for  $D = 1$  and  $n \geq D$ .

Now, let  $D > 1$  and suppose the conclusion of Lemma 11 holds for  $D-1$ . Let  $(a,b) \in R_{(1,1)}^{(\alpha,\beta)}(D)$  and  $n \geq D$ . If  $(a,b) \in R_{(1,1)}^{(\alpha,\beta)}(D-1)$ , then by the induction hypothesis, we're done. Thus, below, we assume  $(a,b) \in R_{(1,1)}^{(\alpha,\beta)}(D) \setminus R_{(1,1)}^{(\alpha,\beta)}(D-1)$ , that is,  $|a-1| + |b-1| = D-1$ . Our goal as before is to show (7).

First, consider the case that  $a > 1$  and  $b > 1$ . Then  $(a-1,b), (a,b-1) \in R_{(1,1)}^{(\alpha,\beta)}(D-1)$  and so by the induction hypothesis  $m_{\mathcal{E}(a-1,b)}^{n-1} = m_{\mathcal{N}(a,b-1)}^{n-1} = 1$  and so (7) holds.

Next, suppose  $a = 1$ . Then  $|b-1| = D-1 > 0$  implies that  $b > 1$ . From the boundary condition (8), we have  $m_{\mathcal{E}(a-1,b)}^{n-1} = 1$ . On the other hand,  $(a,b-1) \in R_{(1,1)}^{(\alpha,\beta)}(D-1)$  and so  $m_{\mathcal{N}(a,b-1)}^{n-1} = 1$ , which proves (7).

The case when  $b = 1$  is analogous to the above argument. This proves the induction step and the lemma.  $\square$

The following definition builds on the notion of the FC property defined previously.

**Definition 12 (Backward Convergence (BC) Property).** Two compatible tuples  $(i_1 \triangleright i_2, \mathcal{D}_1, \mathcal{D}_2)$  and  $(i_3 \triangleright i_4, \mathcal{D}_3, \mathcal{D}_4)$  are said to be *backward convergent* to  $\sigma \in \pm 1$  at time  $n_0 \in \mathbf{N}_0$ , abbreviated as  $\text{BC}(\sigma, n_0)$ , if both tuples are  $\text{FC}(\sigma, n_0)$  and there is exactly one direction in common, i.e.,  $|\{\mathcal{D}_1, \mathcal{D}_2\} \cap \{\mathcal{D}_3, \mathcal{D}_4\}| = 1$ . The unique element  $\mathcal{D}$  not in  $\{\mathcal{D}_1, \mathcal{D}_2\} \cup \{\mathcal{D}_3, \mathcal{D}_4\}$  is called the *backward convergence (BC) direction*.

**Lemma 13 (Backward Convergence).** Let  $\sigma = \pm 1$  and  $n_0 \in \mathbf{N}_0$ . Suppose  $(i_1 \triangleright i_2, \mathcal{D}_1, \mathcal{D}_2)$  and  $(i_3 \triangleright i_4, \mathcal{D}_3, \mathcal{D}_4)$  are  $\text{BC}(\sigma, n_0)$  and let  $\mathcal{D}$  be the BC direction. Then all messages sent from  $R_{i_1}^{i_2} \cap R_{i_3}^{i_4}$  in the direction  $\mathcal{D}$  converge in  $(n_0 + 2N)$  iterations to  $\sigma$ .

*Proof.* Let  $i_\ell = (\alpha_\ell, \beta_\ell)$  for  $\ell = 1, 2, 3, 4$ . Without the loss of generality, we consider the case when  $\mathcal{D}_1 = \mathcal{N}, \mathcal{D}_2 = \mathcal{D}_4 = \mathcal{E}$  and  $\mathcal{D}_3 = \mathcal{S}$ . Hence,  $\mathcal{D} = \mathcal{W}$ .

Furthermore, we assume  $\sigma = 1$  and  $n_0 = 0$ . Now, let  $(a, b) \in R_{(\alpha_1, \beta_1)}^{(\alpha_2, \beta_2)} \cap R_{(\alpha_3, \beta_3)}^{(\alpha_4, \beta_4)}$ . Let  $n \geq 2N$ . Our goal is to show

$$m_{\mathcal{W}(a,b)}^n := \text{sign}\{m_{\mathcal{N}(a,b-1)}^{n-1} + m_{\mathcal{S}(a,b+1)}^{n-1} + m_{\mathcal{W}(a+1,b)}^{n-1}\} = 1 \quad (9)$$

We first show  $m_{\mathcal{N}(a,b-1)}^{n-1} = 1$ . Now, if  $(a, b-1) \in R_{(\alpha_1, \beta_1)}^{(\alpha_2, \beta_2)}$ , then  $m_{\mathcal{N}(a,b-1)}^{n-1} = 1$  by FC Lemma 10. On the other hand, if  $(a, b-1) \notin R_{(\alpha_1, \beta_1)}^{(\alpha_2, \beta_2)}$ , then  $\min(\beta_1, \beta_2) = b$ . Since  $(\mathcal{E}, \mathcal{N})$  is compatible with  $((\alpha_1, \beta_1), (\alpha_2, \beta_2))$  by assumption, we have  $\beta_1 = \min(\beta_1, \beta_2)$ . This shows  $(a, b) = (a, \beta_1) \in L_{(\alpha_1, \beta_1)}^{(\alpha_2, \beta_2)}$ . Next, since  $(a, b-1) + v(\mathcal{N}) = (a, b) \in R_{(\alpha_1, \beta_1)}^{(\alpha_2, \beta_2)}$ ,  $m_{\mathcal{N}(a,b-1)}^n$  is a message received by  $L_{(\alpha_1, \beta_1)}^{(\alpha_2, \beta_2)}$  from the direction  $\mathcal{N}$ . Thus, the defining property of  $((\alpha_1, \beta_1) \triangleright (\alpha_2, \beta_2), \mathcal{N}, \mathcal{E})$  being FC(1, 0) to obtain  $m_{\mathcal{N}(a,b-1)}^{n-1} = 1$ .

An analogous argument shows  $m_{\mathcal{S}(a,b+1)}^{n-1} = 1$ . This proves (9).  $\square$

See Figure 3 for an illustration of Lemma 10 and 13.

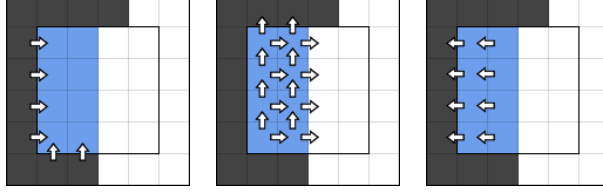


Figure 3: Left: messages received by  $R_{(1,1)}^{(2,4)}$  from the direction  $\mathcal{N}$  and  $\mathcal{E}$ . Middle: convergent messages due to Lemma 10. Right: convergent messages due to Lemma 13.

## 5 Proof of Theorem 6

Let  $\mathbf{x}_{\partial B}$  be a given one-run boundary. Below, we assume the algorithm has run for  $2N$  iterations, i.e.,  $n \geq 2N$ , so that we can use the FC and BC Lemma. The goal is to show that  $\hat{o}_{(a,b)}^n = o_{(a,b)}^*$  for any  $(a, b) \in B$  where  $o_{(a,b)}^*$  is as in Proposition 4 and

$$\hat{o}_{(a,b)}^n = m_{\mathcal{E}(a-1,b)}^n + m_{\mathcal{N}(a,b-1)}^n + m_{\mathcal{W}(a+1,b)}^n + m_{\mathcal{S}(a,b+1)}^n.$$

Without the loss of generality, we assume  $\mathbf{x}_{\partial B}$  satisfies:

1. *Positive run is contracted:* Suppose  $i, j, k \in \partial B$  are such that  $j$  has degree 2,  $i$  and  $k$  are the two neighbors of  $j$ , and  $x_i \neq x_k$ . In such cases, we always assume  $x_j = -1$ . This is because  $j$  does not touch any nodes in  $B$ , so the value of  $x_j$  does not affect the message-updates.

2. *Positive run is smaller* i.e.,  $|R^+| \leq |R^-|$ .

A node in  $B$  with two neighbors in  $\partial B$  is called a *corner*. For easier visualization, the four corners are given names:  $sw = (1, 1)$ ,  $se = (N, 1)$ ,  $ne = (N, N)$

and  $nw = (1, N)$ . Define  $C$  to be the set of corners with two positive boundary neighbors, i.e.,

$$C = \{i \in \{sw, se, ne, nw\} : |\{j \in \partial B \cap \partial i \mid x_j = +1\}| = 2\}$$

Notice that  $|C| \leq 2$  because otherwise  $|R^+| > |R^-|$ . Thus,  $|C| \in \{0, 1, 2\}$  and the proof is correspondingly divided into the three cases.

Below, for brevity, we will write subsets of  $B$  using probabilist notation, i.e., for a logical statement  $S$ , let  $\{S\}$  be a shorthand for  $\{(a, b) \in B : S\}$ . For example,  $\{(a, b) \in B : a = 1\}$  is simply written as  $\{a = 1\}$ .

Case  $|C| = 0$ . Let  $(\alpha_1, \beta_1), (\alpha_2, \beta_2)$  be the two end points of  $R^+$  such that  $\beta_1 \leq \beta_2$ . Without loss of generality, we assume all the positive boundary conditions are restricted to the left side. More precisely, if  $(a, b) \in \partial B$  and  $x_{(a,b)} = 1$ , then  $a = 0$ . See Figure 4 for an example.

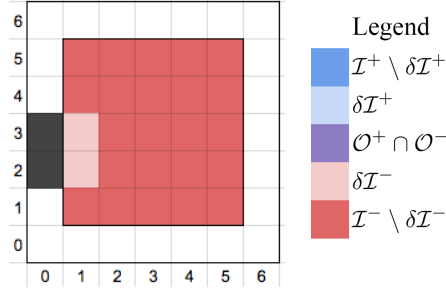


Figure 4: Example of case  $C = \emptyset$  where  $(\alpha_1, \beta_1) = (0, 2), (\alpha_2, \beta_2) = (0, 3)$ .

From the definitions, it is easy to see that

$$\begin{aligned} I^+ \setminus \delta I^+ &= \delta I^+ = O^+ \cap O^- = \emptyset \\ \delta I^- &= \{a = 1, \beta_1 \leq b \leq \beta_2\} \\ I^- \setminus \delta I^- &= \{1 < a\} \cup \{b > \beta_2\} \cup \{b < \beta_1\} \end{aligned}$$

Observe that  $(ne \triangleright sw, \mathcal{S}, \mathcal{W})$  and  $(se \triangleright nw, \mathcal{N}, \mathcal{W})$  are  $BC(-1, 0)$  with backward convergence direction  $\mathcal{E}$ . We will refer to this observation as  $\dagger$ . Thus, for all  $(a, b) \in B$ , we have  $m_{\mathcal{N}(a,b-1)}^n = m_{\mathcal{W}(a+1,b)}^n = m_{\mathcal{S}(a,b+1)}^n = -1$ . We only have to analyze  $m_{\mathcal{E}(a-1,b)}^n$ .

Let  $(a, b) \in I^- \setminus \delta I^-$ . If  $1 < a$ , then by  $\dagger$ ,  $m_{\mathcal{E}(a-1,b)}^n = -1$  and so  $\hat{o}_{(a,b)}^n = -4$ . If  $a = 1$ , then  $m_{\mathcal{E}(a-1,b)}^n = -1$  by the boundary condition. So in both cases,  $\hat{o}_{(a,b)}^n = -4$ .

Next, consider  $(a, b) \in \delta I^-$ , then  $m_{\mathcal{E}(a-1,b)}^n = 1$  by the boundary condition. Hence,  $\hat{o}_{(a,b)}^n = -2$ .

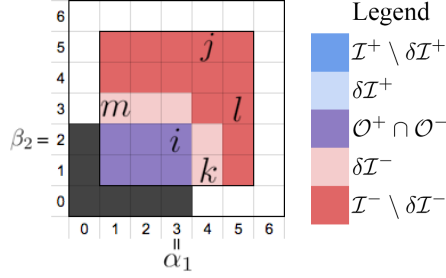


Figure 5:  $|C| = 1$  example,  $(\alpha_1, \beta_1) = (3, 0)$ ,  $(\alpha_2, \beta_2) = (0, 2)$ .

Case  $|C| = 1$ : Without loss of generality, let  $C = \{sw\}$  and let  $(\alpha_1, \beta_1)$  and  $(\alpha_2, \beta_2)$  be the two endpoints of  $R^+$  such that  $\beta_1 \leq \beta_2$ .

One checks that  $\mathcal{I}^+ \setminus \delta \mathcal{I}^+ = \delta \mathcal{I}^+ = \emptyset$  and

$$\begin{aligned} \mathcal{O}^+ \cap \mathcal{O}^- &= \{a \leq \alpha_1, b \leq \beta_2\} \\ \delta \mathcal{I}^- &= \{a = \alpha_1 + 1, b \leq \beta_2\} \cup \{a \leq \alpha_1, b = \beta_2 + 1\} \\ \mathcal{I}^- \setminus \delta \mathcal{I}^- &= \{a > \alpha_1 + 1\} \cup \{b > \beta_2 + 1\} \\ &\quad \cup \{a = \alpha_1 + 1, b = \beta_2 + 1\} \end{aligned}$$

Let  $i = (\alpha_1, \beta_2)$ ,  $j = (\alpha_1 + 1, N)$ ,  $k = (\alpha_1 + 1, 1)$ ,  $l = (N, \beta_2 + 1)$ , and  $m = (1, \beta_2 + 1)$ . See Figure 5. We have the following observations

- I  $(sw \triangleright i, \mathcal{E}, \mathcal{N})$  is FC(1, 0),
- II  $(ne \triangleright sw, \mathcal{W}, \mathcal{S})$  is FC(-1, 0),
- III  $(se \triangleright j, \mathcal{N}, \mathcal{W})$  and  $(ne \triangleright k, \mathcal{S}, \mathcal{W})$  are BC(-1, 0),
- IV  $(nw \triangleright l, \mathcal{E}, \mathcal{S})$  and  $(ne \triangleright m, \mathcal{S}, \mathcal{W})$  are BC(-1, 0).

If  $(a, b) \in \mathcal{O}^+ \cap \mathcal{O}^-$ , then  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{N}(a, b-1)}^n = 1$  by I and  $m_{\mathcal{W}(a+1, b)}^n = m_{\mathcal{S}(a, b+1)}^n = -1$  by II, we have  $\hat{o}_{(a, b)}^n = 0$ .

Let  $(a, b) \in \delta \mathcal{I}^-$ . If  $(a, b) \in \{a = \alpha_1 + 1, b \leq \beta_2\}$ , then  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{N}(a, b-1)}^n = m_{\mathcal{S}(a, b+1)}^n = -1$  by III, and  $m_{\mathcal{W}(a+1, b)}^n = 1$  by I. On the other hand, if  $(a, b) \in \{a \leq \alpha_1, b = \beta_2 + 1\}$ , then  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{N}(a, b-1)}^n = m_{\mathcal{W}(a+1, b)}^n = -1$  by IV, and  $m_{\mathcal{S}(a, b+1)}^n = 1$  by I. Hence, in both cases,  $\hat{o}_{(a, b)}^n = -2$ .

Finally, let  $(a, b) \in \mathcal{I}^- \setminus \delta \mathcal{I}^-$ . If  $(a, b) = (\alpha_1 + 1, \beta_2 + 1)$ , then  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{S}(a, b+1)}^n = -1$  by IV and  $m_{\mathcal{N}(a, b-1)}^n = m_{\mathcal{W}(a+1, b)}^n = -1$  by III. Hence,  $\hat{o}_{(a, b)}^n = -4$ . If  $b > \beta_2 + 1$ , then every messages received by  $(a, b)$  is equal to  $-1$  by IV, so  $\hat{o}_{(a, b)}^n = -4$ . Likewise, if  $a > \alpha_1 + 1$ , then every messages received by  $(a, b)$  is equal to  $-1$  by III, so  $\hat{o}_{(a, b)}^n = -4$ .

Case  $|C| = 2$ : without loss of generality, let  $C = \{sw, nw\}$  and let  $C = \{sw\}$  and let  $(\alpha_1, \beta_1)$  and  $(\alpha_2, \beta_2)$  be the two endpoints of  $R^+$  such that  $\alpha_1 \leq \alpha_2$ .

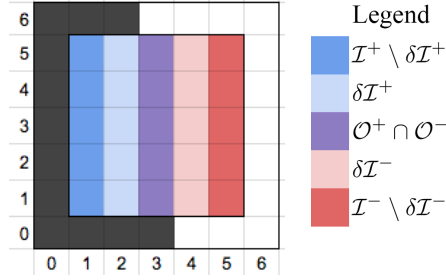


Figure 6:  $|C| = 2$  example where  $(\alpha_1, \beta_1) = (2, 6), (\alpha_2, \beta_2) = (3, 0)$ .

Without the loss of generality, we can further assume  $\alpha_1 \geq \alpha_2$ .  
One checks that

$$\begin{aligned}\mathcal{I}^+ \setminus \delta\mathcal{I}^+ &= \{a < \alpha_1\} \\ \delta\mathcal{I}^+ &= \{a = \alpha_1\} \\ \mathcal{O}^+ \cap \mathcal{O}^- &= \{\alpha_1 < a \leq \alpha_2\} \\ \delta\mathcal{I}^- &= \{a = \alpha_2 + 1\} \\ \mathcal{I}^- \setminus \delta\mathcal{I}^- &= \{\alpha_2 + 1 < a\}\end{aligned}$$

See Figure 6 for an example. Let  $i = (\alpha_2, N)$ ,  $j = (\alpha_1 + 1, 1)$ ,  $k_1 = (\alpha_1, N)$ ,  $k_2 = (\alpha_1, 1)$ ,  $l_1 = (\alpha_2 + 1, 1)$ , and  $l_2 = (\alpha_2 + 1, N)$ . We have the following observations

- i  $(sw \triangleright i, \mathcal{E}, \mathcal{N})$  is  $\text{FC}(1, 0)$ ,
- ii  $(ne \triangleright j, \mathcal{W}, \mathcal{S})$  is  $\text{FC}(-1, 0)$ ,
- iii  $(sw \triangleright k_1, \mathcal{S}, \mathcal{E})$  and  $(nw, k_2, \mathcal{N}, \mathcal{E})$  are  $\text{BC}(-1, 0)$  with BC direction  $\mathcal{W}$ ,
- iv  $(ne \triangleright l_1, \mathcal{W}, \mathcal{S})$  and  $(se, l_2, \mathcal{N}, \mathcal{W})$  are  $\text{BC}(-1, 0)$  with BC direction  $\mathcal{E}$ .

If  $(a, b) \in \mathcal{I}^+ \setminus \delta\mathcal{I}^+$ , then by iii, then  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{N}(a, b-1)}^n = m_{\mathcal{W}(a+1, b)}^n = m_{\mathcal{S}(a, b+1)}^n = 1$  and so  $\hat{o}_{(a, b)}^n = 4$ .

If  $(a, b) \in \delta\mathcal{I}^+$ , then by iii, we have  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{N}(a, b-1)}^n = m_{\mathcal{S}(a, b+1)}^n = 1$  and by ii,  $m_{\mathcal{W}(a+1, b)}^n = -1$ . Hence,  $\hat{o}_{(a, b)}^n = 2$ .

If  $(a, b) \in \mathcal{O}^+ \cap \mathcal{O}^-$ , then by i,  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{N}(a, b-1)}^n = 1$  and by ii,  $m_{\mathcal{W}(a+1, b)}^n = m_{\mathcal{S}(a, b+1)}^n = -1$ . Hence,  $\hat{o}_{(a, b)}^n = 0$ .

If  $(a, b) \in \delta\mathcal{I}^-$ , then by iv,  $m_{\mathcal{N}(a, b-1)}^n = m_{\mathcal{W}(a+1, b)}^n = m_{\mathcal{S}(a, b+1)}^n = -1$  and by i,  $m_{\mathcal{E}(a-1, b)}^n = 1$ . Hence,  $\hat{o}_{(a, b)}^n = -2$ .

If  $(a, b) \in \mathcal{I}^- \setminus \delta\mathcal{I}^-$ , then by iv,  $m_{\mathcal{E}(a-1, b)}^n = m_{\mathcal{N}(a, b-1)}^n = m_{\mathcal{W}(a+1, b)}^n = m_{\mathcal{S}(a, b+1)}^n = -1$  and so  $\hat{o}_{(a, b)}^n = -4$ .

## References

- [1] Y. Weiss, “Correctness of local probability propagation in graphical models with loops,” *Neural computation*, vol. 12, no. 1, pp. 1–41, 2000.
- [2] Y. Weiss and W. T. Freeman, “On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 736–744, 2001.
- [3] M. Wainwright, T. Jaakkola, and A. Willsky, “Tree consistency and bounds on the performance of the max-product algorithm and its generalizations,” *Statistics and computing*, vol. 14, no. 2, pp. 143–166, 2004.
- [4] M. Bayati, D. Shah, and M. Sharma, “Max-product for maximum weight matching: Convergence, correctness, and lp duality,” *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 1241–1251, 2008.
- [5] R. J. Baxter, *Exactly solved models in statistical mechanics*. Courier Corporation, 2007.
- [6] M. G. Reyes, D. L. Neuhoff, and T. N. Pappas, “Lossy cutset coding of bilevel images based on markov random fields,” *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1652–1665, 2014.
- [7] A. Farmer, A. Josan, M. A. Prelee, D. L. Neuhoff, and T. N. Pappas, “Cutset sampling and reconstruction of images,” in *2011 18th IEEE International Conference on Image Processing*. IEEE, 2011, pp. 1909–1912.
- [8] M. G. Reyes, “Cutset based processing and compression of markov random fields,” Ph.D. dissertation, The University of Michigan, 2011.
- [9] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014.